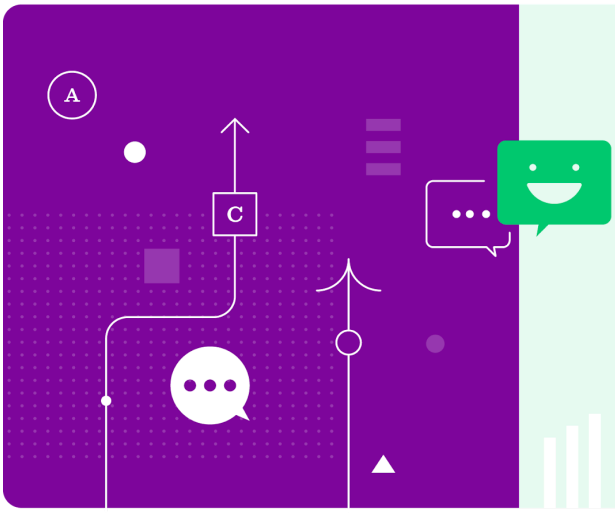
AMIRA LEARNING

Evaluating the Impact of Structured Literacy Implementation on Reading Outcomes in New Mexico

Amira Learning | Every Child Deserves the Chance to Become a Reader
5214F Diamond Heights Blvd, #3255, San Francisco, CA 94131
wwww.AmiraLearning.com | 866-883-7323 | info@amiralearning.com

AMIRA LEARNING



THE INTELLIGENT
Growth Engine

Evaluating the Impact of Structured Literacy Implementation on Reading Outcomes in New Mexico

PREPARED BY
Matthew Jeans, PhD

DATE
June 2025

Executive Summary

This study evaluated the impact of Structured Literacy (SL) implementation on student English and Spanish literacy outcomes across programs in New Mexico elementary schools. Using a quasi-experimental design and propensity score matching, the analysis compared outcomes for students in SL and non-SL schools, collectively drawing on data from nearly 16,000 students across more than 200 schools.

Key Findings

- **Dosage**: Students in SL schools who used Amira more frequently exhibited stronger gains. In particular, students in Grades 1 and 2 (English) and Grades K–2 (Spanish) showed statistically significant moderation effects, suggesting that higher Amira usage in SL schools was associated with improved reading outcomes. For example, an additional 15 minutes per week over 20 weeks translated to a 3-point percentile gain.

- **Performance Level Movement (Spanish)**: In SL schools, students in Grades 3–5 were 61% more likely to move up at least one performance level by the end of the school year compared to peers in non-SL schools. This is a meaningful gain, particularly for ELLs, who comprised most of the Spanish sample.

- **Targeted Benefits for At-Risk Students (English)**: For students beginning the year at the lowest performance levels (Levels 1–2), SL was associated with stronger growth.

- **Language Context May Influence Impact**: The Spanish and English samples had notably different linguistic compositions. The Spanish group was primarily ELLs, whereas the English group was mostly native speakers. SL's structured, explicit nature may have provided greater benefits for ELLs navigating dual-language instruction, contributing to stronger Spanish results.

Structured Literacy, especially when implemented with adequate usage and fidelity, shows promise for improving reading outcomes—particularly for early grade learners and English Learners. While some findings did not reach statistical significance, the direction and consistency of results across models support continued investment in SL-based programs and structured implementation strategies.

Table of Contents

Introduction

Structured Literacy (SL) is a comprehensive and systematic approach to reading instruction grounded in decades of research from cognitive science and linguistics. SL emphasizes explicit instruction in foundational literacy components such as phonology, orthography, morphology, syntax, and semantics, building students' capacity to decode, comprehend, and engage with increasingly complex texts. According to the *New Mexico Public Education Department Biliteracy Guidance* (2022), SL is especially valuable for multilingual learners, offering a linguistically responsive instructional model that supports both English and Spanish literacy development.

Unlike more implicit or unstructured literacy approaches, Structured Literacy provides a carefully sequenced curriculum designed to meet the needs of all learners, including those with or at risk for reading difficulties. Instruction is cumulative and scaffolded, focusing on mastering earlier skills before moving on to more advanced tasks. This makes SL particularly effective for English Language Learners (ELLs), who benefit from explicit connections between spoken and written language and from structured opportunities to develop vocabulary and comprehension in both languages.

In the current study, SL implementation extended beyond instructional content and included a more prescriptive structure for integrating Amira Learning products into classroom practice. SL schools were provided with clear expectations around student usage of the Amira platform, including guidance on frequency, duration, and fidelity of implementation. As a result, SL schools often differed from non-SL schools not only in pedagogy but also in consistency and depth of engagement with the digital assessment and practice tools used throughout the year.

Given these structural and instructional differences, this study sought to evaluate whether SL schools demonstrated different patterns of student growth compared to non-SL schools. Specifically, the inclusion of Amira's usage metrics and assessment outcomes allowed for a closer look at how implementation fidelity, measured in part by time spent on the platform, influenced student performance. These factors are critical in understanding the full impact of SL, as dosage and instructional context can moderate the relationship between program exposure and academic growth.

To that end, this evaluation focused on a large sample of elementary students (K–5) across New Mexico who were assessed in either English (ISIP Assess) or Spanish

(*Evaluar*) and followed throughout the school year. The study addressed the following research questions:

1. What is the impact of SL on student reading achievement in English and Spanish, as measured by percentile growth from beginning-of-year to end-of-year assessments?
2. Does the effect of SL vary as a function of student usage, indicating a potential moderating role of instructional dosage?
3. Among students who began the year with the lowest performance levels, does SL support greater reading growth compared to non-SL campuses?
4. To what extent does SL affect movement across performance levels?

Methodology

Study Overview

This study evaluated the efficacy of a SL intervention implemented in schools across the state of New Mexico in tandem with Amira Instruct (English) and Amira *Ensenar* (Spanish). The study compared student performance in schools using the SL framework against those using a less prescriptive approach to literacy instruction. Analyses were conducted separately for Amira Instruct and Amira *Ensenar*. The outcome measure was performance on the Amira ISIP Assess (English) and Amira ISIP *Evaluar* (Spanish). The goal was to determine whether SL schools demonstrated improved student outcomes in reading achievement and to explore the role of instructional dosage (usage) in student performance.

Analytic Sample

The English sample included 13,244 students from 144 schools, while the Spanish sample included 2,870 students from 62 schools. Students spanned kindergarten through fifth grade. All students were administered benchmark assessments at the beginning (BOY), middle (MOY), and end (EOY) of the academic year.

Propensity score matching was used to create comparable groups of students in SL and non-SL schools. Students were matched on race/ethnicity, gender, grade level, English proficiency, and BOY percentile rankings. Matching was conducted separately for the English and Spanish samples. Students with zero minutes of usage were excluded from the analysis. Matching was conducted without replacement.

Measures

Amira ISIP Assess and Instruct

Formerly known as legacy Istation's ISIP Reading, Amira ISIP Assess is a computer-adaptive testing system designed to monitor students' continuous progress in reading from prekindergarten through 8th grade. It offers a user-friendly experience for teachers and students, with minimal administrative effort required from educators and engaging, developmentally appropriate interfaces for learners. Amira Instruct is a complementary component offering supplementary instructional materials to help students target key areas of improvement for reading intervention.

Based on the science of reading, ISIP Assess and Instruct comprehensively cover essential domains, including phonemic awareness, reading comprehension, listening comprehension, letter knowledge, alphabetic decoding, fluency, spelling, and vocabulary. The assessment's computer-adaptive nature, powered by Item Response Theory (IRT), tailors the difficulty of questions based on each student's performance, ensuring a highly personalized assessment experience (Mathes et al., 2022). This approach enhances the accuracy of measuring student abilities and provides real-time, easily interpretable web-based reports. These reports detail students' strengths and weaknesses, enabling teachers to make informed decisions for targeted instruction and intervention.

Additionally, Amira includes reporting features that automatically alert teachers to students requiring additional instructional support and offers access to a comprehensive library of instructional materials and lessons. This feature aids in customizing teaching strategies to meet individual student needs, which allows for a more accurate profile of each student's abilities while facilitating enhanced teacher planning and student learning outcomes.

Amira ISIP *Evaluar* and *Ensenar*

Formerly known as legacy Istation's *Lectura*, the Amira ISIP *Evaluar* and *Ensenar* are Amira's Spanish literacy assessment and supplementary instruction, respectively. They were not translated versions of Amira's ISIP Assess and Instruct; they were built with respect to the linguistic structure, orthographic transparency, and instructional standards of Spanish. Development began with a Spanish-specific blueprint: Texas Essential Knowledge and Skills (TEKS) for Spanish Language Arts and Reading (SLAR) were merged with standards from California, Puerto Rico, WIDA, and several Latin-American countries to create a single set of grade-by-grade expectations for bilingual classrooms (Istation, 2016, 2019). A national advisory council of researchers in

bilingual education reviewed these standards, and a Texas-based editorial firm (Tri-Lin) authored more than 5,000 items directly in Spanish. No items were translated or adapted from English forms, and every item underwent bias and age-appropriateness reviews (Istation 2016 ,2019).

Amira *Evaluar* targets the domains most predictive of Spanish reading success (i.e., phonemic awareness, grapheme-phoneme conversion, vocabulary, comprehension, and text fluency) that parallel both the Early Grade Reading Assessment (EGRA) used in Latin America and cross-linguistic research on transparent orthographies (Istation 2016 ,2019). Amira *Ensenar* is a supplementary component that complements *Evaluar* to provide targeted reading intervention for students.

Performance Levels and Usage

Both Amira ISIP Assess and Amira *Evaluar* produce performance levels based on percentile rankings. These percentile rankings are grouped into quintiles (i.e., Level 1 through 5), with Level 1 indicating students with the greatest need for intervention. In addition to assessment outcomes, usage data from Amira Instruct and Amira *Ensenar* were recorded as total minutes of engagement from the 2024-25 school year. School-level usage averages were also computed.

Analytical Approach

Data were reshaped from wide to long format to accommodate time-based analyses. Variables representing student-level and school-level usage were mean-centered, and school-level usage was computed as the average usage within each school.

To estimate the effect of the SL intervention on student performance, difference-in-differences (DiD) models were specified using mixed-effects linear regression. These models included fixed effects for SL status, time, and their interaction, with random intercepts at the school level. For models using long-form data, random slopes for time were also tested. The DiD approach allowed estimation of differential changes in reading performance from BOY to EOY, controlling for baseline scores. Sequential models were used to progressively build complexity: initial models included only main effects, followed by interactions and the addition of random slopes and school-level usage. Generally, models were estimated for each grade level in the English sample. For the Spanish sample, models were conducted by grade bands (i.e., K-2 and 3-5) in order to maintain model complexity and increase power due to lower sample sizes.

To assess whether usage moderated the relationship between SL participation and growth, additional multilevel models were specified. These included fixed effects for student-level usage, school-level average usage, and their interactions with SL status and time. All usage variables were centered to aid interpretation and reduce multicollinearity.

For the Spanish sample, an additional analysis was conducted to evaluate movement in performance levels over time. Multilevel ordinal logistic regression (*meologit*) was employed to model changes in performance levels across benchmarks. These models also tested interactions between SL status, time, and usage.

In contrast, for the English sample, further analyses focused on a subsample of students who began the year in performance levels 1 or 2. Because multilevel null models yielded low ICCs in upper grades, linear regression with cluster-robust standard errors was used for this subsample. DiD models and moderation by usage analyses were repeated to assess differential growth among the most at-risk students.

Model fit for all analyses was evaluated using AIC and BIC values, with improvements of 10 points or more considered meaningful. The final model for each analysis was selected based on a combination of statistical fit and interpretability. All mixed-effects models were estimated using Restricted Maximum Likelihood (REML) with Satterthwaite approximations for degrees of freedom to improve the accuracy of fixed effect standard errors and *p*-values, particularly in the presence of unbalanced data or smaller cluster sizes. All data management and statistical analyses were conducted using Stata 19.5.

Results

Amira ISIP Assess and Instruct

Demographic and Descriptive Statistics

Table 1 shows the demographic characteristics of the samples by campus type. The sample is predominately Hispanic, followed by White, American Indian or Alaska Native, Black or African American, Asian, and Other race or ethnicity. Most students are native English speakers, with nearly one-third being English Language Learners (ELL). The sample in SL campuses had more Black or African American as well as

students in other races/ethnicities than the sample in non-SL campuses (3.2 vs. 1.9 and 0.9 vs 0.2, respectively; $p < 0.05$).

Table 1. *Demographic Characteristics of the Sample by Campus Type*

Characteristic	Non-SL	SL
Gender (M)	52.1%	51.4%
Race and Ethnicity		
Hispanic	78.3%	77.2%
White	14.2%	13.5%
American Indian or Alaska Native	4.3%	4.4%
Black or African American	1.9%	3.2%*
Asian	1.2%	0.9%
Other	0.2%	0.9%*
English Proficiency		
Native English Speaker	68.7%	68.4%
English Language Learner (ELL)	31.2%	31.5%
Former ELL	0.1%	0.1%
Sample Size	6622	6622

* $p < 0.05$

Table 2 shows the BOY, MOY, and EOY percentile rank means by grade level and campus type. To examine differences between SL and non-SL schools, ANOVA models were conducted by grade level. All MOY and EOY models controlled for BOY percentile scores. There were no statistically significant differences in percentile ranks at BOY. However, at MOY, non-SL students outperformed their SL peers in Kindergarten (41.2 vs. 39.6; $p < 0.05$), while SL students outperformed their non-SL peers in Grade 3 (43.3 vs. 42.2; $p < 0.05$). By EOY, the performance gap widened, with SL students scoring significantly higher than non-SL students in Grade 3 (44.0 vs. 42.3; $p < 0.05$). In Grade 5, non-SL students outperformed SL students by EOY (42.4 vs. 38.3; $p < 0.05$). In terms of BOY-to-EOY growth, SL students exhibited greater percentile gains in Grade 3 (7.2 vs. 5.2; $p < 0.05$), whereas non-SL students exhibited greater percentile growth in Grade 5 (10.6 vs 8.0; $p < 0.05$) than SL students.

Table 2. *BOY, MOY, and EOY Percentile Rank Means by Grade Level and Campus Type*

Grade	BOY Percentile		MOY Percentile		EOY Percentile		BOY-to-EOY Percentile Gain	
	Non-SL	SL	Non-SL	SL	Non-SL	SL	Non-SL	SL
Kindergarten	33.2 (26.1)	33.5 (26.3)	41.2* (29.2)	39.6 (28.9)	47.4 (30.8)	46.4 (30.4)	14.3 (24.1)	12.9 (23.1)
Grade 1	34.3 (29.7)	33.9 (29.5)	40.4 (29.5)	41.2 (29.4)	39.2 (29.5)	40.1 (29.9)	4.9 (17.1)	6.2 (20.0)
Grade 2	33.0 (29.3)	33.1 (29.3)	39.6 (31.5)	38.5 (32.0)	41.1 (32.5)	40.4 (32.1)	8.1 (17.2)	7.3 (16.8)
Grade 3	37.1 (31.2)	36.8 (30.9)	42.2 (32.5)	43.3* (32.3)	42.3 (31.8)	44.0* (31.7)	5.2 (14.5)	7.2* (15.7)
Grade 4	33.6 (28.3)	32.3 (28.3)	40.4 (30.6)	38.5 (30.1)	41.4 (31.1)	39.5 (29.7)	7.8 (16.1)	7.2 (17.1)
Grade 5	31.8 (29.5)	30.3 (28.9)	39.8 (32.3)	37.1 (30.5)	42.4* (33.8)	38.3 (30.8)	10.6* (18.9)	8.0 (15.0)
Total Sample	34.0 (29.1)	33.6 (29.0)	40.7 (31.1)	39.9 (30.6)	42.3 (31.6)	41.8 (30.9)	8.3 (18.5)	8.2 (18.5)

* $p < 0.05$

Table 3 presents the proportion of total students in each performance level at each benchmark by campus type. Grade-specific tables can be found in Appendix A. Performance level distributions revealed broadly similar patterns between SL and non-SL schools across the year, with most changes reflecting modest shifts. At BOY, 43.5% of students in non-SL schools and 44.2% in SL schools were in Level 1. By EOY, this proportion declined to 34.2% and 34.0%, respectively, indicating a 9–10% reduction in Level 1 representation for both groups. Gains in higher performance levels were distributed somewhat evenly across Levels 3 through 5, though non-SL schools showed slightly more students in Level 5 at EOY compared to SL schools (17.2% vs. 15.4%).

Table 3. *Proportion of Total Students in Performance Levels at BOY, MOY, and EOY Benchmarks by Campus Type*

Benchmark	Level 1		Level 2		Level 3		Level 4		Level 5	
	Non-SL	SL	Non-SL	SL	Non-SL	SL	Non-SL	SL	Non-SL	SL
BOY	43.5%	44.2%	19.8%	19.3%	15.0%	15.0%	12.0%	11.9%	9.8%	9.6%
N	2880	2927	1308	1281	995	994	791	786	648	634
MOY	35.5%	36.5%	19.3%	18.6%	14.9%	15.8%	14.9%	14.8%	15.5%	14.3%
N	2351	2418	1277	1232	986	1046	985	981	1023	945
EOY	34.2%	34.0%	18.2%	18.6%	15.2%	15.4%	15.2%	16.5%	17.2%	15.4%
N	2266	2253	1207	1233	1004	1022	1005	1095	1140	1019

Table 4 presents the usage as average total minutes for each grade level by campus type for the 2024-25 school year. On average, SL students had significantly higher usage in Grade 1 (+60.4, $p < 0.05$), Grade 2 (+138.9, $p < 0.05$), and overall (+25.6, $p < 0.05$) compared to non-SL students. One exception was observed in Grade 4, where non-SL students had higher usage (+49.7, $p < 0.05$) than non-SL students. Overall, there is an apparent shift in usage decreasing in grades 4 and 5 in SL schools compared to non-SL schools.

Table 4. *Average Usage in Minutes by Grade Level and Campus Type*

Grade	Non-SL	SL
Kindergarten	394.5 (482.4)	384.5 (435.6)
Grade 1	415.0 (443.4)	475.4* (471.6)
Grade 2	378.3 (442.9)	517.2* (517.7)
Grade 3	358.5 (384.9)	369.8 (325.2)
Grade 4	437.5* (368.4)	387.8 (328.7)
Grade 5	341.6 (389.2)	306.2 (333.5)
Total Sample	390.3 (424.6)	415.9* (421.7)

* $p < 0.05$

Amira ISIP Assess Difference-in-Differences Analysis

This portion of the analysis was designed to examine whether students in SL schools experienced greater growth in reading performance across the school year compared to their peers in non-SL schools. By evaluating changes in percentile ranks

from BOY, MOY, and EOY of the academic year, the analysis sought to determine whether SL implementation was associated with stronger academic gains in reading as measured by Amira ISIP Assess.

Although no statistically significant effects were found in any of the grade-level models, Grades 1 and 3 showed some positive trends. In Grade 1, the interaction coefficient between SL status and the MOY and EOY timepoints was 1.19 and 1.24, respectively, suggesting that students in SL schools had, on average, slightly higher gains than their non-SL peers, though the result did not reach statistical significance. In Grade 3, the corresponding interaction coefficient for MOY and EOY was 1.39 and 2.04, respectively, indicating a modest positive effect of SL on growth in reading scores. While small, these effects reflect potential trends that may warrant further investigation in larger samples or in studies with longer-term follow-up.

Complete model estimates for each grade level can be found in Appendix B.

Moderation by Amira Instruct

This analysis examined whether the impact of the SL framework on student reading growth was moderated by instructional dosage using Amira Instruct, operationalized as student usage minutes.

Among the six grade levels analyzed, statistically significant moderation effects were detected in four grades, shown below in Tables 5–8. However, these results presented mixed results. In Grades 1 and 2, there was a significant positive interaction between SL and usage ($\beta = 0.01$ and $\beta = 0.005$, respectively), indicating that students in SL schools who used the tool more frequently demonstrated greater gains in percentile scores than their peers in non-SL schools.

Table 5. *Model Estimates Examining Usage as a Moderator of Reading Growth in Grade 1 Students by Campus Type*

Fixed Effects	Model 1	Model 2	Model 3	Model 4
Intercept	36.34* (1.31)	32.38* (1.60)	32.38* (1.60)	32.22* (1.62)
Structured Literacy (SL)		-0.62 (3.05)	-0.62 (3.05)	-0.45 (3.07)
Centered Usage		0.01* (0.001)	0.01* (0.001)	0.01* (0.001)
SL x Centered Usage		0.01* (0.002)	0.01* (0.002)	0.01* (0.002)
MOY		6.71* (0.77)	6.71* (0.77)	6.71* (0.77)
EOY		5.56* (0.77)	5.56* (0.77)	5.56* (0.77)
Error Variance				
Level-1	716.28* (12.19)	689.98* (11.75)	689.98* (11.75)	689.95* (11.75)
Level-2 Intercept	189.01* (27.65)	198.34* (29.05)	198.34* (29.05)	199.47* (29.27)
Time			1.11e-11* (3.26e-09)	7.65e-11* (1.88e-08)
Centered Average School Usage				-0.005 (0.01)
Model Fit				
AIC	66491.71	66259.63	66261.63	66271.34
BIC	66512.28	66314.5	66323.36	66339.92

* $p < 0.05$; ICC = .21

Table 6. *Model Estimates Examining Usage as a Moderator of Reading Growth in Grade 2 Students by Campus Type*

Fixed Effects	Model 1	Model 2	Model 3	Model 4
Intercept	36.84* (1.33)	33.46* (1.64)	33.46* (1.64)	33.26* (1.65)
Structured Literacy (SL)		-4.07 (3.15)	-4.07 (3.15)	-3.84 (3.16)
Centered Usage		0.01* (0.001)	0.01* (0.001)	0.01* (0.002)
SL x Centered Usage		0.005* (0.002)	0.005* (0.002)	0.005* (0.002)
MOY		6.01* (0.80)	6.01* (0.80)	6.01* (0.80)
EOY		7.70* (0.80)	7.70* (0.80)	7.70* (0.80)
Error Variance				
Level-1	837.98* (13.74)	815.31* (13.38)	815.31* (13.38)	815.27* (13.38)
Level-2 Intercept	196.21* (28.93)	205.25* (30.35)	205.25* (30.38)	205.55* (30.50)
Time			3.85e-08* (7.49e-06)	6.49e-07* (0.0001)
Centered Average School Usage				-0.01 (0.01)
Model Fit				
AIC	72701.64	72523.05	72525.05	72534.19
BIC	72722.43	72578.5	72587.43	72603.5

* $p < 0.05$; ICC = .19

Although the coefficients appear small, their cumulative effect over time can be meaningful. For instance, in Grade 1, a student using the tool for an additional 15 minutes per week over 20 weeks (300 minutes total) would be expected to gain approximately 3 percentile points more than a comparable student in a non-SL school. Similarly, in Grade 2, the same usage pattern would translate to an estimated 1.5 percentile point advantage. These findings suggest that consistent usage in SL schools may yield modest but potentially important gains, particularly in the early grades where reading development is most sensitive to instructional input.

However, the pattern was reversed in upper elementary. In Grade 3, there was a significant negative interaction between SL and usage (β = –0.01). A similar negative interaction was observed in Grade 5 (β = –0.01), suggesting that in those grades, increased usage in SL schools was associated with slightly lower gains relative to non-SL schools with similar usage levels. The results for Grades K and 4 were not statistically significant and can be found in Appendix C.

Table 7. *Model Estimates Examining Usage as a Moderator of Reading Growth in Grade 3 Students by Campus Type*

Fixed Effects	Model 1	Model 2	Model 3	Model 4
Intercept	39.26* (1.36)	35.37* (1.63)	35.37* (1.63)	35.37* (1.64)
Structured Literacy (SL)		-0.25 (3.12)	-0.25 (3.12)	-0.26 (3.14)
Centered Usage		0.01* (0.002)	0.01* (0.002)	0.01* (0.002)
SL x Centered Usage		-0.01* (0.003)	-0.01* (0.003)	-0.01* (0.003)
MOY		5.79* (0.83)	5.79* (0.83)	5.79* (0.83)
EOY		6.19* (0.83)	6.19* (0.83)	6.19* (0.83)
Error Variance				
Level-1	862.92* (14.37)	852.08* (14.20)	852.08* (14.20)	852.08* (14.20)
Level-2 Intercept	194.63* (30.04)	191.81* (30.01)	191.80* (30.01)	193.68* (30.43)
Time			2.34e-09* (7.76e-07)	2.28e-09*
Centered Average School Usage				0.0001 (0.01)
Model Fit				
AIC	70776.74	70703.5	70705.5	70713.94
BIC	70797.44	70758.72	70767.62	70776.05

* $p < 0.05$; ICC = .18

Table 8. *Model Estimates Examining Usage as a Moderator of Reading Growth in Grade 5 Students by Campus Type*

Fixed Effects	Model 1	Model 2	Model 3	Model 4
Intercept	34.70* (1.67)	29.91* (2.39)	29.97* (2.35)	29.47* (2.38)
Structured Literacy (SL)		-2.98 (3.46)	-2.81 (3.40)	-2.42 (3.38)
Centered Usage		-0.002 (0.002)	-0.002 (0.003)	-0.003 (0.003)
SL x Centered Usage		-0.01* (0.004)	-0.01* (0.004)	-0.01* (0.004)
MOY		7.36* (1.13)	7.24* (1.16)	7.24* (1.16)
EOY		9.27* (1.13)	9.03* (1.27)	9.04* (1.27)
Error Variance				
Level-1	859.61* (19.55)	835.80* (19.02)	833.80* (19.02)	833.95* (19.03)
Level-2 Intercept	152.02 (34.37)	163.69* (36.69)	154.38* (36.23)	150.98* (36.17)
Time			3.52* (3.18)	3.54* (3.18)
Centered Average School Usage				0.01 (0.01)
Model Fit				
AIC	37962.56	37874.03	37874.23	37882.93
BIC	37981.39	37924.27	37930.74	37945.73

* $p < 0.05$; ICC = .15

These findings indicate that instructional dosage played a differential role in the effectiveness of the SL framework, with positive moderation effects in the early grades (1–2) and negative or null effects in upper grades (3–5). These contrasting patterns highlight the importance of considering developmental stage when evaluating program impact and dosage thresholds.

Difference-in-Differences Analysis Among Level 1 and 2 Students

Following the full-sample DiD and moderation analyses, a secondary analysis was conducted on a high-priority subgroup: students who scored in Levels 1 or 2 at BOY. These students are considered most in need of intervention. The goal of this analysis was to determine whether SL schools demonstrated stronger growth for these at-risk learners compared to their peers in non-SL schools and whether usage of the supplemental tool moderated these effects.

Among the six grade levels analyzed, statistically significant effects were found in Grade 3. As shown in Table 9, the interaction between SL status and the MOY and EOY benchmarks was statistically significant, with coefficients of 2.08 and 2.74,

respectively. These results suggest that among students beginning the year in Level 1 or 2, those in SL schools demonstrated greater gains in percentile scores over time compared to their peers in non-SL schools.

Table 9. *Difference-in-Differences Model Estimates of Reading Growth in Grade 3 At-Risk Students by Campus Type*

Fixed Effects	Model 1	Model 2	Model 3
Intercept	19.42* (0.60)	14.16* (0.46)	14.24* (0.48)
Structured Literacy (SL)		-0.03 (0.82)	-0.11 (0.82)
MOY		6.16* (0.59)	6.16* (0.59)
EOY		7.23* (0.61)	7.24* (0.61)
SL x MOY		2.08* (1.04)	2.08* (1.04)
SL x EOY		2.74* (1.13)	2.74* (1.14)
Centered Average School Usage			0.002 (0.003)
Model Fit			
AIC	37395.4	37201.73	37200.84
BIC	37401.77	37239.96	37245.46

* $p < 0.05$

Although other grade-level models did not yield statistically significant interactions, Grade 1 showed a notable positive trend, shown in Table D2. In this model, the interaction terms for SL status with MOY and EOY were 1.74 and 1.21, respectively. While these coefficients did not reach significance, they suggest a consistent directional trend of improved performance for SL students over the course of the academic year. Results for all other grade levels can be found in Appendix D.

Moderation by Amira Instruct for Level 1 and 2 Students

To further examine whether the benefits of SL varied by instructional dosage, a moderation analysis was conducted among students who scored in Levels 1 or 2 at BOY.

Among all grade levels, Grade 4 yielded the strongest evidence of a statistically significant moderation effect, shown in Table 10. The interaction term between SL status and centered usage was positive and statistically significant ($\beta = 0.01$). This result indicates that for Level 1 and 2 students in Grade 4, increased usage of the instructional tool was more strongly associated with reading gains in SL schools than in non-SL schools. Translating this into dosage terms, an additional 15 minutes per week over a 20-week span (i.e., 300 minutes total) predicted an approximate 3-point gain in percentile score, suggesting a meaningful impact for at-risk students.

Table 10. *Model Estimates Examining Usage as a Moderator of Reading Growth in Grade 4 At-Risk Students by Campus Type*

Fixed Effects	Model 1	Model 2	Model 3
Intercept	19.97* (0.67)	14.38* (0.78)	14.06* (0.85)
Structured Literacy (SL)		-0.36 (1.32)	-0.22 (1.31)
Centered Usage		0.001 (0.002)	-0.001 (0.002)
SL x Centered Usage		0.01 (0.003)	0.01* (0.003)
MOY		7.88* (0.54)	7.88* (0.54)
EOY		9.64* (0.88)	9.64* (0.88)
Centered Average School Usage			0.003 (0.003)
Model Fit			
AIC	36708.98	36481.24	36478.64
BIC	36715.32	36519.32	36523.07

* $p < 0.05$

Although not statistically significant, the coefficients in Grades K–3 were also positive, suggesting a consistent trend across lower grades. In Kindergarten and Grade 1, the interaction was β = 0.002. In practical terms, 15 extra minutes per week over 20 weeks (300 minutes total) would yield a gain of 0.6 percentile points in SL schools compared to non-SL schools. Grade 2 showed a slightly lower interaction coefficient of β = 0.001, suggesting a gain of 0.3 percentile points for equivalent dosage. Grade 3 exhibited the same β = 0.01 value seen in Grade 4 but was not statistically significant. The tables for all other grade levels can be found in Appendix E.

These findings, while not uniformly significant, point to a potentially important role for dosage in supporting the reading growth of struggling students in SL settings. They suggest that further investigation, perhaps with a larger sample or over longer periods, may be warranted to more definitively assess the benefits of sustained usage in these contexts.

Amira *Evaluar* and *Ensenar*

Demographic and Descriptive Statistics

Table 11 shows the demographic characteristics of the samples by campus type. The sample is predominately Hispanic, followed by White, and Other race or ethnicity. Most students are ELLs, with approximately one-third being native English speakers. There were no statistically significant differences in these demographic characteristics between SL and non-SL schools.

Table 11. *Demographic Characteristics of the Sample by Campus Type*

Characteristic	Non-SL	SL
Gender (F)	50.5%	50.3%
Race and Ethnicity		
Hispanic	95.2%	95.5%
White	4.1%	3.9%
Other	0.7%	0.6%
English Proficiency		
Native English Speaker	33.4%	32.6%
English Language Learner (ELL)	66.6%	67.4%
Sampel Size	1435	1435

* $p < 0.05$

Table 12 shows the BOY, MOY, and EOY percentile rank means by grade band and campus type. To examine differences between SL and non-SL schools, ANOVA models were conducted by grade band. All MOY and EOY models controlled for BOY percentile scores. There were no statistically significant differences in percentile ranks at BOY. However, at MOY, SL students outperformed their non-SL peers in Grades 3–5 (31.4 vs. 28.4; $p < 0.05$) and overall (35.8 vs. 33.1; $p < 0.05$). By EOY, the performance gap widened, with SL students scoring significantly higher than non-SL students in Grades K–2 (39.4 vs. 33.9; $p < 0.05$), Grades 3–5 (33.9 vs. 28.6; $p < 0.05$), and overall (37.7 vs. 32.3; $p < 0.05$).

In terms of growth from BOY to EOY, students in SL schools demonstrated significantly greater gains than their non-SL peers in Grades 3–5, with an average percentile increase of 8.8 points compared to just 2.4 points in non-SL schools ($p <$ .05). In contrast, students in Grades K–2 showed average declines in percentile scores across both groups. However, the decline was less pronounced in SL schools (-4.4) than in non-SL schools (-8.9), suggesting a potential buffering effect of SL instruction. When considering the full sample, both groups exhibited small declines in growth,

but the decrease was more modest in SL settings compared to non-SL settings (-0.5 vs. -5.6).

Table 12. *BOY, MOY, and EOY Percentile Rank Means by Grade Band and Campus Type*

Grade Band	BOY Percentile		MOY Percentile		EOY Percentile		BOY-to-EOY Percentile Gain	
	Non-SL	SL	Non-SL	SL	Non-SL	SL	Non-SL	SL
Grades K-2 (N=2018)	42.8 (25.4)	43.7 (25.8)	35.1 (27.3)	37.6 (29.2)	33.9 (27.5)	39.4* (30.3)	-8.9 (25.1)	-4.4* (27.3)
Grades 3-5 (N=852)	26.2 (22.3)	25.1 (21.4)	28.4 (26.6)	31.4* (27.9)	28.6 (27.9)	33.9* (29.7)	2.4 (16.1)	8.8* (18.8)
Total Sample (N=2870)	37.9 (25.6)	38.2 (26.0)	33.1 (27.3)	35.8* (28.9)	32.3 (27.7)	37.7* (30.2)	-5.6 (23.4)	-0.5* (25.8)

* $p < 0.05$

Table 13 presents the proportion of total students in each performance level at each benchmark by campus type. Grade band-specific tables can be found in Appendix F. At BOY, the percentage of students in Level 1 was similar across both groups (30.5% in non-SL vs. 31.4% in SL schools). However, by MOY and EOY, the proportion of students in Level 1 increased more sharply in non-SL schools, reaching 44.0% at EOY compared to 36.9% in SL schools, suggesting that fewer SL students remained in the lowest performance band over time. Correspondingly, slightly more SL students reached Levels 4 and 5 by the end of the year (11.2% in Level 4 and 13.9% in Level 5) than students in non-SL schools (10.7% and 8.1%, respectively), indicating a modest but notable upward trend in performance distribution in favor of SL campuses.

Table 13. *Proportion of Total Students in Performance Levels at BOY, MOY, and EOY Benchmarks by Campus Type*

Benchmark	Level 1		Level 2		Level 3		Level 4		Level 5	
	Non-SL	SL	Non-SL	SL	Non-SL	SL	Non-SL	SL	Non-SL	SL
BOY	30.5%	31.4%	27.8%	25.6%	20.0%	19.9%	13.5%	14.5%	8.2%	8.4%
N	438	450	399	371	287	286	194	208	117	120
MOY	41.3%	39.3%	23.3%	21.7%	16.2%	16.2%	11.4%	12.3%	7.8%	10.5%
N	592	564	335	312	232	233	164	176	112	150
EOY	44.0%	36.9%	21.8%	21.5%	15.4%	16.5%	10.7%	11.2%	8.1%	13.9%
N	632	530	313	309	221	236	153	160	116	200

Table 14 presents the usage as average total minutes for each grade band by campus type for the 2024-25 school year. On average, SL students had significantly higher usage in Grades K–2 (+37.3, $p < 0.05$), Grades 3–5 (+124.0, $p < 0.05$), and overall (+63.1, $p < 0.05$) compared to non-SL students.

Table 14. *Average Usage in Minutes by Grade Band and Campus Type*

Grade	Non-SL	SL
Grades K-2	289.8 (298.3)	327.1* (334.4)
Grades 3-5	121.8 (166.6)	245.8* (228.5)
Total Sample	239.9 (276.9)	303.0* (309.0)

* $p < 0.05$

Amira *Evaluar* Difference-in-Differences Analysis

This portion of the analysis was designed to examine whether students in SL schools experienced greater growth in Spanish literacy performance across the school year compared to their peers in non-SL schools. By evaluating changes in percentile ranks from BOY, MOY, and EOY of the academic year, the analysis sought to determine whether SL implementation was associated with stronger academic gains in reading as measured by Amira *Evaluar*.

For early elementary students (K–2), a significant interaction was observed in Model 2 ($\beta = 4.52$), shown in Table 15. This suggests that students in SL schools demonstrated greater growth from BOY to EOY compared to their peers in non-SL schools. However, this interaction term became non-significant in Model 3, which accounted for random slopes for time and demonstrated better model fit. The lack of statistical significance in the final model implies that the observed growth advantage for SL students may not be robust once after accounting for variability across schools in growth trajectories. This pattern indicates a potential signal of impact that warrants further exploration, particularly as gains appeared more pronounced by the end of the school year.

Table 15. *Difference-in-Differences Model Estimates of Spanish Literacy Growth in Grades K–2 by Campus Type*

Fixed Effects	Model 1	Model 2	Model 3	Model 4
Intercept	36.38* (1.53)	41.39* (2.10)	41.98* (1.76)	41.72* (1.64)
Structured Literacy (SL)		-0.75 (3.29)	-0.37 (2.72)	-1.51 (2.55)
MOY		-7.71* (1.16)	-8.37* (1.41)	-8.25* (1.40)
EOY		-8.90* (1.16)	-10.21* (2.01)	-9.99* (1.98)
SL x MOY		1.62 (1.64)	1.18 (2.07)	1.12 (2.05)
SL x EOY		4.52* (1.64)	3.63 (3.04)	3.53 (2.99)
Error Variance				
Level-1	685.25* (12.50)	674.53* (12.31)	667.40* (12.20)	667.48* (12.20)
Level-2 Intercept	88.12* (21.40)	90.58* (22.19)	54.99* (15.64)	45.05* (13.42)
Time			13.95* (5.06)	13.28* (4.81)
Centered Average School Usage				0.02* (0.01)
Model Fit				
AIC	56826.19	56725.29	56687.97	56690.1
BIC	56846.31	56778.96	56748.34	56757.18

* $p < 0.05$; ICC = .11

For upper elementary students, a statistically significant interaction between SL status and the EOY timepoint was observed in Model 2 ($\beta = 6.41$, $p < .05$), shown in Table 16. This result suggests that students in SL schools made greater gains in Spanish literacy from BOY to EOY compared to their non-SL peers. Although Models 3 and 4 introduced additional complexity (e.g., random slopes and school-level usage), they did not lead to appreciable improvements in model fit. Therefore, Model 2 was interpreted. These findings indicate a meaningful positive impact of the Structured Literacy approach on Spanish literacy outcomes for students in Grades 3–5.

Table 16. *Difference-in-Differences Model Estimates of Spanish Literacy Growth in Grades 3–5 by Campus Type*

Fixed Effects	Model 1	Model 2	Model 3	Model 4
Intercept	26.10* (1.74)	23.42* (2.41)	24.02* (2.25)	24.27* (2.22)
Structured Literacy (SL)		-0.38 (3.83)	-0.34 (3.55)	-2.06 (3.60)
MOY		2.10 (1.63)	1.51 (1.76)	1.55 (1.76)
EOY		2.35 (1.63)	1.16 (2.11)	1.25 (2.10)
SL x MOY		4.14 (2.31)	4.08 (2.54)	4.10 (2.53)
SL x EOY		6.41* (2.31)	6.28 (3.14)	6.33 (3.11)
Error Variance				
Level-1	575.95* (16.25)	569.39* (16.08)	565.68* (16.02)	565.60* (16.01)
Level-2 Intercept	117.15* (28.96)	117.99* (29.43)	92.97*(26.51)	88.28* (25.20)
Time			8.09* (4.72)	7.75* (4.54)
Centered Average School Usage				
Model Fit				
AIC	23599.23	23560.56	23555.68	23561.35
BIC	23616.77	23607.33	23608.29	23619.81

* $p < 0.05$; ICC = .17

Moderation by Amira *Ensenar*

This analysis examined whether the impact of the SL framework on student Spanish literacy growth was moderated by instructional dosage using Amira *Ensenar*, operationalized as student usage minutes.

In the K–2 grade band, a statistically significant positive interaction between SL and centered usage was observed, shown in Table 17. The coefficient is $\beta = 0.01$, indicating that greater usage was associated with larger gains in percentile scores for students in SL schools compared to their peers in non-SL schools. Translating this to a more tangible dosage effect: an increase of 15 minutes of weekly use over 20 weeks (i.e., 300 additional minutes) would correspond to an approximate 3-point increase in percentile rank.

Table 17. *Model Estimates Examining Usage as a Moderator of Spanish Literacy Growth in Grades K–2 by Campus Type*

Fixed Effects	Model 1	Model 2	Model 3	Model 4
Intercept	36.38* (1.53)	40.33* (1.97)	41.73* (1.64)	41.46* (1.58)
Structured Literacy (SL)		1.43 (3.02)	0.19 (2.49)	-0.70 (2.40)
Centered Usage		-0.001 (0.002)	-0.001 (0.002)	-0.002 (0.002)
SL x Centered Usage		0.01* (0.003)	0.01* (0.002)	0.01* (0.002)
MOY		-6.90* (0.82)	-7.91* (1.03)	-7.82* (1.02)
EOY		-6.64* (0.82)	-8.65* (1.51)	-8.48* (1.48)
Error Variance				
Level-1	685.25* (12.50)	672.33* (12.27)	664.49* (12.15)	664.56* (12.15)
Level-2 Intercept	88.12* (21.40)	82.20* (20.36)	48.33* (14.09)	43.15* (12.87)
Time			14.03* (4.96)	13.31* (12.87)
Centered Average School Usage				0.02* (0.01)
Model Fit				
AIC	56826.19	56729.19	56686.21	56690.81
BIC	56846.31	56782.86	56746.59	56757.89

* $p < 0.05$; ICC = .11

In contrast, results for Grades 3–5 showed a positive, but non-significant, interaction between SL and usage, shown in Table 18. While the direction of the effect was consistent with that observed in K–2, the lack of statistical significance suggests weaker or more variable effects in upper elementary Spanish literacy instruction.

Table 18. *Model Estimates Examining Usage as a Moderator of Spanish Literacy Growth in Grades 3–5 by Campus Type*

Fixed Effects	Model 1	Model 2	Model 3	Model 4
Intercept	26.10* (1.74)	21.29* (2.38)	22.29* (2.22)	22.80* (2.19)
Structured Literacy (SL)		3.54 (3.67)	2.34 (3.38)	0.60 (3.43)
Centered Usage		-0.004 (0.005)	-0.004 (0.005)	-0.005 (0.005)
SL x Centered Usage		0.002 (0.01)	0.002 (0.01)	0.001 (0.01)
MOY		4.17* (1.16)	3.33* (1.29)	3.39* (1.28)
EOY		5.56* (1.16)	3.86* (1.62)	3.99* (1.61)
Error Variance				
Level-1	575.95* (16.25)	570.67* (16.12)	565.64* (16.03)	565.49* (16.02)
Level-2 Intercept	117.15* (28.96)	122.11* (30.66)	94.66* (27.09)	89.89* (25.59)
Time			9.70* (5.09)	9.31* (4.91)
Centered Average School Usage				0.02* (0.01)
Model Fit				
AIC	23599.23	23592.08	23584.07	23589.32
BIC	23616.77	23638.85	23636.68	23647.79

* $p < 0.05$; ICC = .17

Performance Level Movement

Given the descriptive analyses showing shifts in student distributions across performance levels by benchmark period (BOY, MOY, and EOY), a secondary analysis was conducted to evaluate whether students in SL schools were more likely to move into higher performance levels over time. This investigation was especially pertinent in the Spanish sample, where performance level changes were visibly more pronounced than raw percentile gains.

For the K–2 grade band, the interaction between SL status and the EOY benchmark period was statistically significant ($\beta = 0.27$), shown in Table 19. This suggests that students in SL schools were more likely to move into higher performance levels by the end of the year. This coefficient translates to a 31% increase in the odds of moving up a level for SL students compared to non-SL peers.

Table 19. *Multilevel Ordinal Logistic Regression Results Predicting Performance Level Movement in Grades K–2 by Campus Type*

Fixed Effects	Model 1	Model 2	Model 3	Model 4
Structured Literacy (SL)		-0.03 (0.22)	-0.01 (0.17)	-0.07 (0.16)
MOY		-0.53* (0.08)	-0.59* (0.10)	-0.58* (0.10)
EOY		-0.63* (0.08)	-0.76* (0.15)	-0.74* (0.15)
SL x MOY		0.08 (0.11)	0.06 (0.15)	0.06 (0.15)
SL x EOY		0.32* (0.11)	0.27 (0.22)	0.27 (0.22)
Centered Usage			-0.0001 (0.0002)	-0.0001 (0.0002)
SL x Centered Usage			0.001 (0.0003)	0.001 (0.0003)
MOY x Centered Usage			0.0001 (0.0003)	0.0001 (0.0003)
EOY x Centered Usage			0.0001 (0.0003)	0.0002 (0.0003)
SL x MOY x Centered Usage			0.0001 (0.0004)	0.0001 (0.0004)
SL x EOY x Centered Usage			0.0003 (0.0004)	0.0003 (0.0004)
Threshold 1	-0.64* (0.10)	-0.99* (0.14)	-1.05* (0.11)	-0.76* (0.16)
Threshold 2	0.47* (0.10)	0.14 (0.14)	0.09)0.11)	0.38* (0.16)
Threshold 3	1.38* (0.11)	1.07* (0.14)	1.02* (0.11)	1.31* (0.16)
Threshold 4	2.43* (0.11)	2.11* (0.15)	2.07* (0.12)	2.37* (0.16)
Error Variance				
Level-2 Intercept	0.39* (0.10)	0.40* (0.10)	0.21* (0.06)	0.18* (0.05)
Time			0.08* (0.03)	0.07* (0.03)
Centered Average School Usage				0.001* (0.001)
Model Fit				
AIC	18009.72	17915.37	17856.14	17852.26
BIC	18043.26	17982.45	17970.19	17973.01

* $p < 0.05$; ICC = .11

For Grades 3–5, the interaction between SL status and EOY was again statistically significant (β = 0.47), as shown in Table 20. This corresponds to a 61% increase in the odds of achieving a higher performance level for SL students at EOY compared to students in non-SL schools.

Table 20. *Multilevel Ordinal Logistic Regression Results Predicting Performance Level Movement in Grades 3–5 by Campus Type*

Fixed Effects	Model 1	Model 2	Model 3	Model 4
Structured Literacy (SL)		-0.05 (0.31)	0.01 (0.31)	-0.11 (0.31)
MOY		0.17 (0.13)	0.17 (0.15)	0.17 (0.15)
EOY		0.12 (0.13)	0.12 (0.17)	0.12 (0.17)
SL x MOY		0.25 (0.19)	0.22 (0.20)	0.22 (0.20)
SL x EOY		0.47* (0.19)	0.37 (0.22)	0.37 (0.22)
Centered Usage			-0.0005 (0.001)	-0.0005 (0.0006)
SL x Centered Usage			0.0001 (0.001)	0.00002 (0.001)
MOY x Centered Usage			0.0004 (0.001)	0.0004 (0.001)
EOY x Centered Usage			0.001 (0.001)	0.001 (0.001)
SL x MOY x Centered Usage			-0.0004 (0.001)	-0.0004 (0.001)
SL x EOY x Centered Usage			0.0001 (0.001)	0.0001 (0.001)
Threshold 1	0.13 (0.14)		0.31 (0.20)	0.56* (0.26)
Threshold 2	1.16* (0.14)		1.35* (0.20)	1.60* (0.26)
Threshold 3	2.09* (0.15)		2.29* (0.21)	2.54* (0.26)
Threshold 4	3.11* (0.16)		3.31* (0.21)	3.57* (0.27)
Error Variance				
Level-2 Intercept	0.75* (0.19)	0.76* (0.19)	0.73* (0.20)	0.69* (0.19)
Time			0.01* (0.02)	0.01* (0.02)
Centered Average School Usage				0.001 (0.001)
Model Fit				
AIC	6580.859	6567.223	6577.067	6576.714
BIC	6610.09	6625.69	6676.452	6681.945

* $p < 0.05$; ICC = .19

No significant moderation by usage was observed in either grade band. These findings suggest that, particularly by EOY, students in SL schools were more likely to experience positive shifts in their performance level classification compared to students in non-SL schools, especially in the upper elementary grades. Although moderation by usage did not emerge as a significant factor, the EOY benchmark period appears to be a critical point for detecting program impacts on student performance categorization.

Discussion

This study examined the impact of SL implementation on student reading outcomes across both English and Spanish literacy in New Mexico elementary schools. Using a robust quasi-experimental design and rigorous statistical modeling, the study provides valuable insights into how SL may influence student performance, particularly when dosage and grade-level considerations are taken into account. While not all findings were statistically significant, positive trends consistently emerged, especially for younger students and those with higher engagement.

Although the main effects of SL on percentile growth were not statistically significant across all models, promising trends were observed in multiple grades. For example, Grades 1 and 3 in the English sample showed positive coefficients, suggesting that students in SL schools were growing at a faster rate than their non-SL peers, even if the differences did not reach conventional thresholds of statistical significance. These trends align with the developmental importance of early reading interventions and underscore SL's potential to catalyze growth in lower elementary grades.

The most consistent evidence for SL's positive impact emerged in models that accounted for instructional dosage. For Amira Instruct, significant moderation effects were observed in Grades 1 and 2, and in Amira *Ensenar*, the K–2 grade band demonstrated a statistically significant interaction between SL and student-level usage. In these models, students in SL schools who engaged more frequently with the literacy tool experienced greater gains than similarly engaged peers in non-SL schools. These results affirm that SL is most effective when students have adequate exposure and suggest that dosage is a critical lever for realizing program benefits.

Complementary analyses focusing on performance level movement offered additional support for SL's impact, particularly for Amira *Evaluar*. Among Grade 3–5 students in SL schools, the likelihood of moving up at least one performance level was 61% higher than for non-SL peers, even after controlling for usage. This finding is particularly encouraging given the high percentage of ELL students in the Amira *Evaluar* sample, reinforcing the value of a structured, linguistically responsive approach. These categorical improvements may better capture meaningful progress among students with initially low proficiency, especially when percentile gains are more difficult to detect.

Subgroup analyses among students in performance Levels 1 or 2 at BOY, the most at-risk learners, revealed some promising patterns. While effects were not

consistently significant across all grades, Grade 3 students in SL schools demonstrated statistically significant growth from BOY to EOY in English. In addition, moderation models for these at-risk students indicated positive, though not always significant, trends in Grades K–3, culminating in a significant effect in Grade 4. These findings highlight the potential of SL to benefit students with the greatest need when applied consistently and supported by sufficient instructional time.

One important factor in interpreting these results is the demographic and linguistic context of the two samples. While race and ethnicity were well-matched across SL and non-SL groups, the linguistic profiles of the Amira ISIP Assess and Amira *Evaluar* samples differed notably. The ISIP Assess sample was composed primarily of native English speakers, with approximately one-third classified as ELLs. In contrast, the *Evaluar* sample was predominantly composed of ELLs, with one-third identified as native Spanish speakers. These distinct linguistic contexts may help explain the variation in results across languages. For example, SL may provide essential scaffolding for Spanish-speaking ELLs navigating dual-language instruction, thereby enhancing its impact in Spanish settings.

Together, the findings from this study suggest that SL can support improved student outcomes, especially when implemented early and accompanied by sufficient instructional dosage. While some effects did not reach statistical significance, the consistent direction of findings, especially in K–2 grades and among high-usage students, reinforces the value of SL for building early reading proficiency. The results also underscore the importance of monitoring and supporting implementation fidelity, especially in linguistically diverse settings. Future research should continue to explore how dosage, language status, and instructional context interact to influence outcomes on state assessments, such as the New Mexico Measures of Student Success and Achievement, and should aim to capture longitudinal trends that may further clarify SL's impact over time.

References

Istation. (2016). *ISIP Español technical manual: Computer adaptive testing system for continuous progress monitoring of reading growth for students Pre-K through Grade 3*.

Istation. (2019). *ISIP Lectura Avanzada technical manual: Computer adaptive testing system for continuous progress monitoring of Spanish reading abilities for students in Grades 4 and 5*.

Mathes, P., Torgesen, J., & Herron, J. (2022). Istation's Indicators of Progress Technical Report: Computer Adaptive System for Continuous Progress Monitoring of Reading Growth for Students Pre-K through Grade 8: Istation.

New Mexico Public Education Department, & Cavazos, L. (2022). *Biliteracy guidance: The science of reading for English learners*. https://webnew.ped.state.nm.us/wp-content/uploads/2022/12/NMPED-Biliteracy-Guidance_12.4.22.pdf

Appendix A

Table A1. *Proportion of Kindergarten Students in Performance Levels at BOY, MOY, and EOY Benchmarks by Campus Type*

Benchmark	Level 1		Level 2		Level 3		Level 4		Level 5	
	Non-SL	SL	Non-SL	SL	Non-SL	SL	Non-SL	SL	Non-SL	SL
BOY	40.1%	40.7%	25.2%	23.4%	16.5%	17.1%	12.0%	12.6%	6.2%	6.2%
N	481	489	303	281	198	205	144	151	75	75
MOY	32.1%	34.3%	21.9%	21.5%	17.7%	18.1%	14.8%	13.7%	13.5%	12.5%
N	386	412	263	258	212	217	178	164	162	150
EOY	27.6%	28.8%	17.8%	16.5%	15.7%	15.8%	19.0%	22.3%	19.9%	16.6%
N	332	346	214	198	188	190	228	268	239	199

Table A2. *Proportion of Grade 1 Students in Performance Levels at BOY, MOY, and EOY Benchmarks by Campus Type*

Benchmark	Level 1		Level 2		Level 3		Level 4		Level 5	
	Non-SL	SL	Non-SL	SL	Non-SL	SL	Non-SL	SL	Non-SL	SL
BOY	44.9%	45.0%	19.5%	19.9%	13.1%	13.1%	11.2%	11.6%	11.4%	10.5%
N	526	527	228	233	153	153	131	136	134	123
MOY	34.0%	32.1%	22.8%	22.2%	14.7%	17.9%	14.9%	13.6%	13.7%	14.3%
N	398	376	267	260	172	210	174	159	161	167
EOY	37.5%	35.1%	18.9%	21.3%	15.8%	14.5%	15.0%	14.7%	12.8%	14.4%
N	439	411	222	250	185	170	176	172	150	169

Table A3. *Proportion of Grade 2 Students in Performance Levels at BOY, MOY, and EOY Benchmarks by Campus Type*

Benchmark	Level 1		Level 2		Level 3		Level 4		Level 5	
	Non-SL	SL	Non-SL	SL	Non-SL	SL	Non-SL	SL	Non-SL	SL
BOY	44.7%	44.7%	18.4%	17.7%	15.4%	15.7%	12.3%	12.6%	9.2%	9.3%
N	564	564	232	223	194	198	155	159	116	117
MOY	40.9%	43.1%	13.9%	13.2%	14.5%	12.9%	14.4%	16.9%	16.3%	14.0%
N	516	544	175	166	183	162	181	213	206	176
EOY	36.8%	37.8%	16.8%	16.4%	15.6%	14.8%	12.9%	14.8%	17.8%	16.3%
N	464	477	212	207	197	186	163	186	225	205

Table A4. *Proportion of Grade 3 Students in Performance Levels at BOY, MOY, and EOY Benchmarks by Campus Type*

Benchmark	Level 1		Level 2		Level 3		Level 4		Level 5	
	Non-SL	SL	Non-SL	SL	Non-SL	SL	Non-SL	SL	Non-SL	SL
BOY	41.8%	42.0%	17.2%	17.0%	15.4%	15.9%	11.5%	11.6%	14.2%	13.6%
N	511	514	210	208	188	194	141	142	174	166
MOY	34.2%	32.2%	18.1%	19.0%	13.6%	15.2%	16.0%	15.1%	18.0%	18.6%
N	419	394	222	232	167	186	196	185	220	227
EOY	32.5%	30.5%	19.9%	19.0%	16.5%	16.6%	12.8%	15.4%	18.4%	18.6%
N	398	373	243	232	202	203	156	188	225	228

Table A5. *Proportion of Grade 4 Students in Performance Levels at BOY, MOY, and EOY Benchmarks by Campus Type*

Benchmark	Level 1		Level 2		Level 3		Level 4		Level 5	
	Non-SL	SL	Non-SL	SL	Non-SL	SL	Non-SL	SL	Non-SL	SL
BOY	43.1%	45.1%	19.3%	19.4%	16.7%	15.4%	13.6%	11.8%	7.2%	8.3%
N	477	499	214	215	185	170	151	131	80	92
MOY	34.3%	38.5%	20.1%	18.1%	15.7%	15.6%	15.0%	15.0%	14.8%	12.8%
N	380	426	223	200	174	173	166	166	164	142
EOY	34.7%	34.7%	18.5%	21.3%	15.0%	15.2%	16.9%	17.1%	14.9%	11.7%
N	384	384	205	236	166	168	187	189	165	130

Table A6. *Proportion of Grade 5 Students in Performance Levels at BOY, MOY, and EOY Benchmarks by Campus Type*

Benchmark	Level 1		Level 2		Level 3		Level 4		Level 5	
	Non-SL	SL	Non-SL	SL	Non-SL	SL	Non-SL	SL	Non-SL	SL
BOY	48.9%	50.8%	18.4%	18.4%	11.7%	11.3%	10.5%	10.2%	10.5%	9.3%
N	321	334	121	121	77	74	69	67	69	61
MOY	38.4%	40.5%	19.3%	17.7%	11.9%	14.9%	13.7%	14.3%	16.7%	12.6%
N	252	266	127	116	78	98	90	94	110	83
EOY	37.9%	39.9%	16.9%	16.7%	10.1%	16.0%	14.5%	14.0%	20.7%	13.4%
N	249	262	111	110	66	105	95	92	136	88

Appendix B

Table B1. *Difference-in-Differences Model Estimates of Reading Growth in Kindergarten Students by Campus Type*

Fixed Effects	Model 1	Model 2	Model 3	Model 4
Intercept	38.44* (1.15)	30.53* (1.47)	30.90* (1.39)	30.83* (1.40)
Structured Literacy (SL)		2.87 (2.80)	2.92 (2.60)	3.00 (2.62)
MOY		8.04* (1.07)	7.68* (1.16)	7.68* (1.16)
EOY		14.26* (1.07)	13.53* (1.40)	13.53* (1.41)
SL x MOY		-1.95 (1.51)	-2.04 (1.75)	-2.05 (1.76)
SL x EOY		-1.32 (1.51)	-1.50 (2.34)	-1.51 (2.36)
Error Variance				
Level-1	719.88* (12.10)	688.67* (11.58)	680.01* (11.50)	679.93* (11.50)
Level-2 Intercept	139.95* (20.94)	142.34* (21.23)	118.00* (19.34)	118.94* (19.55)
Time			15.55* (4.52)	15.80* (4.61)
Centered Average School Usage				-0.002 (0.005)
Model Fit				
AIC	68133.73	67814.70	67782.92	67793.54
BIC	68154.38	67869.76	67844.86	67862.37

* $p < 0.05$; ICC = .16

Table B2. *Difference-in-Differences Model Estimates of Reading Growth in Grade 1 Students by Campus Type*

Fixed Effects	Model 1	Model 2	Model 3	Model 4
Intercept	36.34* (1.31)	32.89* (1.65)	32.89* (1.65)	33.10* (1.66)
Structured Literacy (SL)		-1.75 (3.14)	-1.75 (3.14)	-1.97 (3.14)
MOY		6.12* (1.10)	6.12* (1.10)	6.12* (1.10)
EOY		4.94* (1.10)	4.94 (1.10)	4.94 (1.10)
SL x MOY		1.19 (1.55)	1.19 (1.55)	1.19 (1.55)
SL x EOY		1.24 (1.55)	1.24 (1.55)	1.24 (1.55)
Error Variance				
Level-1	716.28* (12.19)	707.85* (12.05)	707.85* (12.05)	707.88* (12.05)
Level-2 Intercept	189.01* (27.65)	191.12* (28.00)	191.12* (28.00)	190.30* (28.04)
Time			5.89e-11* (1.65e-08)	2.46e-12* (7.39e-10)
Centered Average School Usage				0.006 (0.006)
Model Fit				
AIC	66491.71	66404.37	66406.37	66415.62
BIC	66512.28	66459.23	66468.09	66484.21

* $p < 0.05$; ICC = .21

Table B3. *Difference-in-Differences Model Estimates of Reading Growth in Grade 2 Students by Campus Type*

Fixed Effects	Model 1	Model 2	Model 3	Model 4
Intercept	36.84* (1.33)	32.54* (1.68)	32.54* (1.68)	32.67* (1.69)
Structured Literacy (SL)		-1.79 (3.23)	-1.79 (3.23)	-1.95 (3.24)
MOY		6.62* (1.15)	6.62* (1.15)	6.62* (1.15)
EOY		8.12* (1.15)	8.12* (1.15)	8.12* (1.15)
SL x MOY		-1.22 (1.62)	-1.22 (1.62)	-1.22 (1.62)
SL x EOY		-0.84 (1.62)	-0.84 (1.62)	-0.84 (1.62)
Error Variance				
Level-1	837.98* (13.74)	827.22* (13.57)	827.22* (13.57)	827.23* (13.58)
Level-2 Intercept	196.21* (28.93)	197.63* (29.17)	197.63* (29.17)	198.37* (29.49)
Time			0.0001*	8.09e-10* (2.02e-07)
Centered Average School Usage				0.004 (0.006)
Model Fit				
AIC	72701.64	72599.14	72599.14	72599.14
BIC	72722.43	72654.59	72654.59	72654.59

* $p < 0.05$; ICC = .19

Table B4. *Difference-in-Differences Model Estimates of Reading Growth in Grade 3 Students by Campus Type*

Fixed Effects	Model 1	Model 2	Model 3	Model 4
Intercept	39.26* (1.36)	35.86* (1.71)	35.86* (1.71)	36.04* (1.71)
Structured Literacy (SL)		-1.25 (3.30)	-1.25 (3.30)	-1.56 (3.29)
MOY		5.10* (1.18)	5.10* (1.18)	5.10* (1.18)
EOY		5.17* (1.18)	5.17* (1.18)	5.17* (1.18)
SL x MOY		1.39 (1.67)	1.39 (1.67)	1.39 (1.67)
SL x EOY		2.04 (1.67)	2.04 (1.67)	2.04 (1.67)
Error Variance				
Level-1	862.92* (14.37)	855.04* (14.24)	855.04* (14.24)	855.15* (14.24)
Level-2 Intercept	194.63* (30.04)	197.25* (30.50)	197.25* (30.50)	194.00* (30.39)
Time			3.89e-09* (6.91e-07)	4.75e-09* (8.28e-07)
Centered Average School Usage				0.008 (0.006)
Model Fit				
AIC	70776.74	70704.47	70706.47	70715.18
BIC	70797.44	70759.69	70768.59	70784.19

* $p < 0.05$; ICC = .18

Table B5. *Difference-in-Differences Model Estimates of Reading Growth in Grade 4 Students by Campus Type*

Fixed Effects	Model 1	Model 2	Model 3	Model 4
Intercept	33.89* (1.51)	28.26* (2.07)	28.30* (2.07)	27.51* (2.11)
Structured Literacy (SL)		2.24 (3.27)	2.13 (3.26)	2.73 (3.26)
MOY		6.86* (1.19)	6.82* (1.21)	6.82* (1.21)
EOY		7.78* (1.19)	7.70* (1.25)	7.71* (1.25)
SL x MOY		-0.63 (1.69)	-0.52 (1.71)	-0.52 (1.71)
SL x EOY		-0.53 (1.69)	-0.31 (1.78)	-0.31 (1.78)
Error Variance				
Level-1	797.46* (13.94)	786.56* (13.75)	785.98* (13.76)	785.82* (13.76)
Level-2 Intercept	151.89* (31.42)	155.67* (32.18)	154.51* (32.18)	152.55* (31.82)
Time			0.89* (1.46)	0.93* (1.47)
Centered Average School Usage				0.01 (0.01)
Model Fit				
AIC	63418.33	63322.15	63323.69	63331.17
BIC	63438.73	63376.56	63384.90	63399.18

* $p < 0.05$; ICC = .16

Table B6. *Difference-in-Differences Model Estimates of Reading Growth in Grade 5 Students by Campus Type*

Fixed Effects	Model 1	Model 2	Model 3	Model 4
Intercept	34.70* (1.67)	29.45* (2.43)	29.62* (2.39)	29.40* (2.43)
Structured Literacy (SL)		-0.86 (3.62)	-0.95 (3.55)	-0.80 (3.57)
MOY		7.94* (1.60)	7.79* (1.65)	7.80* (1.65)
EOY		10.56* (1.60)	10.26* (1.80)	10.27* (1.80)
SL x MOY		-1.16 (2.27)	-1.08 (2.34)	-1.08 (2.34)
SL x EOY		-2.57 (2.27)	-2.40 (2.54)	-2.40 (2.54)
Error Variance				
Level-1	859.61* (19.55)	843.89* (19.20)	842.02* (19.21)	842.06* (19.21)
Level-2 Intercept	152.02* (34.37)	155.10* (35.24)	145.66* (34.84)	147.29* (35.55)
Time			3.49* (3.25)	3.46* (3.25)
Centered Average School Usage				
Model Fit				
AIC	37962.56	37882.07	37882.39	37892.31
BIC	37981.39	37932.3	37938.90	37955.10

* $p < 0.05$; ICC = .15

Appendix C

Table C1. *Model Estimates Examining Usage as a Moderator of Reading Growth in Kindergarten Students by Campus Type*

Fixed Effects	Model 1	Model 2	Model 3	Model 4
Intercept	38.44* (1.15)	30.95* (1.41)	31.15* (1.37)	30.79* (1.36)
Structured Literacy (SL)		2.16 (2.66)	2.50 (2.55)	2.93 (2.52)
Centered Usage		0.01* (0.001)	0.01* (0.001)	0.01* (0.001)
SL x Centered Usage		-0.001 (0.002)	-0.004 (0.002)	-0.001 (0.002)
MOY		7.07* (0.75)	6.79* (0.85)	6.77* (0.86)
EOY		13.60* (0.75)	13.04* (1.09)	13.01* (1.12)
Error Variance				
Level-1	719.88* (12.10)	684.50* (11.51)	676.32* (11.44)	675.83* (11.43)
Level-2 Intercept	139.95* (20.94)	142.87* (21.29)	123.99* (19.99)	119.43* (19.38)
Time			13.84* (4.18)	15.03* (4.42)
Centered Average School Usage				-0.01* (0.01)
Model Fit				
AIC	68133.73	67799.77	67771.98	67777.15
BIC	68154.38	67854.83	67833.92	67845.98

* $p < 0.05$; ICC = .16

Table C2. *Model Estimates Examining Usage as a Moderator of Reading Growth in Grade 4 Students by Campus Type*

Fixed Effects	Model 1	Model 2	Model 3	Model 4
Intercept	33.89* (1.51)	28.67* (2.01)	28.65* (2.01)	28.03* (2.08)
Structured Literacy (SL)		1.71 (3.11)	1.72 (3.11)	2.20 (3.14)
Centered Usage		0.004* (0.002)	0.004* (0.002)	0.003 (0.002)
SL x Centered Usage		-0.003 (0.003)	-0.003 (0.003)	-0.003 (0.003)
MOY		6.54* (0.84)	6.56* (0.85)	6.56* (0.85)
EOY		7.51* (0.84)	7.55* (0.89)	7.55* (0.89)
Error Variance				
Level-1	797.46* (13.94)	786.09* (13.74)	785.83* (13.75)	785.42* (13.75)
Level-2 Intercept	151.89* (31.42)	154.46* (31.78)	153.33* (31.77)	154.06* (32.00)
Time			0.83* (1.40)	0.86* (1.41)
Centered Average School Usage				0.01 (0.01)
Model Fit				
AIC	63418.33	63344.48	63346.04	63354.86
BIC	63438.73	63398.89	63407.25	63422.87

* $p < 0.05$; ICC = .16

Appendix D

Table D1. *Difference-in-Differences Model Estimates of Reading Growth in Kindergarten At-Risk Students by Campus Type*

Fixed Effects	Model 1	Model 2	Model 3	Model 4
Intercept	26.05* (0.82)	16.57* (1.13)	16.92* (0.74)	17.02* (0.73)
Structured Literacy (SL)		-0.25 (2.05)	-0.43 (1.10)	-0.61 (1.10)
MOY		11.35* (1.00)	10.93* (1.24)	10.94* (1.23)
EOY		18.07* (1.00)	17.23* (1.80)	17.25* (1.78)
SL x MOY		-1.87 (1.43)	-1.87 (2.08)	-1.84 (2.06)
SL x EOY		-1.19 (1.43)	-1.20 (3.39)	-1.13 (3.34)
Error Variance				
Level-1	448.03* (9.41)	395.25* (8.31)	377.33* (7.97)	377.58* (7.98)
Level-2 Intercept	62.27* (10.84)	65.51* (11.11)	4.21* (3.40)	3.77* (3.37)
Time			49.28* (8.63)	47.41* (8.40)
Centered Average School Usage				0.004 (0.002)
Model Fit				
AIC	41889.21	41316.43	41148.95	41157.7
BIC	41908.55	41368.01	41206.97	41222.17

* $p < 0.05$; ICC = .12

Table D2. *Difference-in-Differences Model Estimates of Reading Growth in Grade 1 At-Risk Students by Campus Type*

Fixed Effects	Model 1	Model 2	Model 3	Model 4
Intercept	20.98* (0.68)	15.03* (0.95)	14.98* (0.85)	15.28* (0.85)
Structured Literacy (SL)		-0.90 (1.69)	-0.42 (1.46)	-0.76 (1.43)
MOY		8.87* (0.87)	8.92* (0.92)	8.91* (0.92)
EOY		8.47* (0.87)	8.58* (1.07)	8.55* (1.07)
SL x MOY		2.22 (1.22)	1.74 (1.37)	1.77 (1.37)
SL x EOY		2.18 (1.22)	1.21 (1.75)	1.27 (1.74)
Error Variance				
Level-1	304.45* (6.48)	282.52* (6.02)	279.62* (5.98)	279.70* (5.99)
Level-2 Intercept	42.39* (7.60)	44.27* (7.81)	27.59* (7.15)	25.47* (6.81)
Time			7.27* (2.74)	7.15* (2.70)
Centered Average School Usage				0.01* (0.003)
Model Fit				
AIC	39058.32	38726.47	38706.89	38712.53
BIC	39077.59	38777.84	38764.68	38776.74

* $p < 0.05$; ICC = .12

Table D3. *Difference-in-Differences Model Estimates of Reading Growth in Grade 2 At-Risk Students by Campus Type*

Fixed Effects	Model 1	Model 2	Model 3	Model 4
Intercept	19.36* (0.73)	14.70* (0.99)	14.28* (0.87)	14.49* (0.88)
Structured Literacy (SL)		-2.08 (1.80)	-1.38 (1.52)	-1.64 (1.52)
MOY		6.78* (0.87)	7.12* (0.93)	7.11* (0.92)
EOY		9.31* (0.87)	10.00* (1.10)	9.97* (1.09)
SL x MOY		-0.52 (1.23)	-1.14 (1.40)	-1.11 (1.39)
SL x EOY		-0.87 (1.23)	-2.10 (1.82)	-2.03 (1.79)
Error Variance				
Level-1	314.42* (6.55)	300.01* (6.26)	296.70* (6.21)	296.75* (6.21)
Level-2 Intercept	50.67* (9.14)	51.93* (9.25)	31.32* (8.29)	30.97* (8.02)
Time			8.72* (3.10)	8.13* (2.92)
Centered Average School Usage				0.01* (0.003)
Model Fit				
AIC	41002.92	40783.36	40761.8	40769.11
BIC	41022.32	40835.08	40819.99	40833.77

* $p < 0.05$; ICC = .14

Table D4. *Difference-in-Differences Model Estimates of Reading Growth in Grade 4 At-Risk Students by Campus Type*

Fixed Effects	Model 1	Model 2	Model 3
Intercept	19.97* (0.67)	14.35* (0.50)	13.93* (0.61)
Structured Literacy (SL)		-0.44 (1.23)	-0.22 (1.19)
MOY		8.28* (0.80)	8.28* (0.80)
EOY		9.35* (1.28)	9.35* (1.28)
SL x MOY		-0.78 (1.08)	-0.78 (1.08)
SL x EOY		0.57 (1.76)	0.57 (1.76)
Centered Average School Usage			0.004 (0.002)
Model Fit			
AIC	36708.98	36503.01	36494.26
BIC	36715.32	36541.09	36538.69

* $p < 0.05$

Table D5. *Difference-in-Differences Model Estimates of Reading Growth in Grade 5 At-Risk Students by Campus Type*

Fixed Effects	Model 1	Model 2	Model 3
Intercept	19.69* (0.89)	13.63* (0.64)	13.46* (0.80)
Structured Literacy (SL)		-0.34 (0.98)	-0.20 (1.00)
MOY		8.38* (0.87)	8.38* (0.87)
EOY		11.39* (2.11)	11.39* (2.11)
SL x MOY		-0.07 (1.44)	-0.07 (1.44)
SL x EOY		-2.05 (2.37)	-2.05 (2.37)
Centered Average School Usage			0.001 (0.003)
Model Fit			
AIC	23530.09	23384.73	23385.85
BIC	23535.98	23420.12	23427.13

* $p < 0.05$

Appendix E

Table E1. *Model Estimates Examining Usage as a Moderator of Reading Growth in Kindergarten At-Risk Students by Campus Type*

Fixed Effects	Model 1	Model 2	Model 3	Model 4
Intercept	26.05* (0.82)	16.95* (1.00)	17.07* (0.72)	17.00* (0.72)
Structured Literacy (SL)		-0.90 (1.77)	-0.73 (1.05)	-0.61 (1.05)
Centered Usage		0.01* (0.001)	0.004* (0.001)	0.005* (0.001)
SL x Centered Usage		0.001 (0.002)	0.002 (0.002)	0.002 (0.002)
MOY		10.42* (0.71)	10.17* (0.94)	10.16* (0.94)
EOY		17.47* (0.71)	16.97* (1.45)	16.96* (1.46)
Error Variance				
Level-1	448.03* (9.41)	391.99* (8.24)	374.37* (7.91)	374.25* (7.91)
Level-2 Intercept	62.27* (10.84)	56.77* (9.88)	6.88* (3.82)	6.91* (3.81)
Time			42.56* (7.73)	43.25* (7.86)
Centered Average School Usage				-0.003 (0.003)
Model Fit				
AIC	41889.21	41293.32	41140.08	41151.22
BIC	41908.55	41344.90	41198.11	41215.69

* $p < 0.05$; ICC = .12

Table E2. *Model Estimates Examining Usage as a Moderator of Reading Growth in Grade 1 At-Risk Students by Campus Type*

Fixed Effects	Model 1	Model 2	Model 3	Model 4
Intercept	20.98* (0.68)	14.44* (0.87)	14.77* (0.80)	14.84* (0.81)
Structured Literacy (SL)		0.75 (1.51)	0.43 (1.33)	0.35 (1.34)
Centered Usage		0.005* (0.001)	0.005* (0.001)	0.005* (0.001)
SL x Centered Usage		0.001 (0.001)	0.001 (0.001)	0.002 (0.001)
MOY		9.98* (0.61)	9.73* (0.67)	9.73* (0.67)
EOY		9.56* (0.61)	9.05* (0.83)	9.05* (0.83)
Error Variance				
Level-1	304.45* (6.48)	279.35* (5.95)	276.35* (5.91)	276.41* (5.92)
Level-2 Intercept	42.39* (7.60)	42.60* (7.6)	28.16* (7.01)	28.40* (7.05)
Time			6.66* (2.49)	6.54* (2.46)
Centered Average School Usage				0.002 (0.003)
Model Fit				
AIC	39058.32	38700.69	38681.11	38692.56
BIC	39077.59	38752.06	38738.9	38756.77

* $p < 0.05$; ICC = .12

Table E3. *Model Estimates Examining Usage as a Moderator of Reading Growth in Grade 2 At-Risk Students by Campus Type*

Fixed Effects	Model 1	Model 2	Model 3	Model 4
Intercept	19.36* (0.73)	15.24* (0.93)	14.87* (0.84)	14.93* (0.84)
Structured Literacy (SL)		-3.09 (1.65)	-2.49 (1.42)	-2.57 (1.44)
Centered Usage		0.004* (0.001)	0.003* (0.001)	0.003* (0.001)
SL x Centered Usage		0.001 (0.002)	0.001 (0.001)	0.001 (0.002)
MOY		6.52* (0.61)	6.68* (0.68)	6.68* (0.68)
EOY		8.88* (0.61)	9.21* (0.84)	9.20* (0.85)
Error Variance				
Level-1	314.42* (6.55)	298.12* (6.21)	294.96* (6.18)	294.99* (6.18)
Level-2 Intercept	50.67* (9.14)	51.21* (9.10)	32.54* (8.28)	33.30* (8.36)
Time			7.85* (2.86)	7.59* (2.79)
Centered Average School Usage				0.002 (0.003)
Model Fit				
AIC	41002.92	40780.59	40761.03	40772.37
BIC	41022.32	40832.32	40819.22	40837.03

* $p < 0.05$; ICC = .14

Table E4. *Model Estimates Examining Usage as a Moderator of Reading Growth in Grade 3 At-Risk Students by Campus Type*

Fixed Effects	Model 1	Model 2	Model 3
Intercept	19.42* (0.60)	13.32* (0.61)	13.34* (0.61)
Structured Literacy (SL)		1.62 (1.15)	1.58 (1.14)
Centered Usage		-0.001 (0.002)	-0.002 (0.002)
SL x Centered Usage		0.01 (0.003)	0.01 (0.003)
MOY		7.20* (0.55)	7.20* (0.55)
EOY		8.61* (0.61)	8.61* (0.61)
Centered Average School Usage			0.002 (0.003)
Model Fit			
AIC	37395.4	37189.6	37189.41
BIC	37401.77	37227.84	37234.02

* $p < 0.05$

Table E5. *Model Estimates Examining Usage as a Moderator of Reading Growth in Grade 5 At-Risk Students by Campus Type*

Fixed Effects	Model 1	Model 2	Model 3
Intercept	19.69* (0.89)	13.99* (1.01)	13.71* (1.11)
Structured Literacy (SL)		-1.04 (1.79)	-0.82 (1.77)
Centered Usage		-4.25e-06 (0.003)	-0.001 (0.003)
SL x Centered Usage		-0.001 (0.004)	-0.0003 (0.004)
MOY		8.34* (0.72)	8.34* (0.72)
EOY		10.35* (1.19)	10.35* (1.19)
Centered Average School Usage			0.003 (0.004)
Model Fit			
AIC	23530.09	23385.97	23386.17
BIC	23535.98	23421.36	23427.46

* $p < 0.05$

Appendix F

Table F1. *Proportion of Grade K–2 Students in Performance Levels at BOY, MOY, and EOY Benchmarks by Campus Type*

Benchmark	Level 1		Level 2		Level 3		Level 4		Level 5	
	Non-SL	SL	Non-SL	SL	Non-SL	SL	Non-SL	SL	Non-SL	SL
BOY	23.1%	22.9%	27.6%	25.6%	22.3%	22.5%	17.2%	18.3%	9.8%	10.7%
N	233	231	278	258	225	227	174	185	99	108
MOY	37.8%	36.1%	24.8%	22.9%	16.7%	16.9%	12.6%	12.4%	8.2%	11.8%
N	381	364	250	231	168	170	127	125	83	119
EOY	40.3%	33.9%	24.5%	22.7%	16.2%	17.3%	10.5%	10.9%	8.5%	15.2%
N	407	342	247	229	163	175	106	110	86	153

Table F2. *Proportion of Grade 3–5 Students in Performance Levels at BOY, MOY, and EOY Benchmarks by Campus Type*

Benchmark	Level 1		Level 2		Level 3		Level 4		Level 5	
	Non-SL	SL	Non-SL	SL	Non-SL	SL	Non-SL	SL	Non-SL	SL
BOY	48.1%	51.4%	28.4%	26.5%	14.6%	13.9%	4.7%	5.4%	4.2%	2.8%
N	205	219	121	113	62	59	20	23	18	12
MOY	49.5%	47.0%	20.0%	19.0%	15.0%	14.8%	8.7%	12.0%	6.8%	7.3%
N	211	200	85	81	64	63	37	51	29	31
EOY	52.8%	44.1%	15.5%	18.8%	13.6%	14.3%	11.0%	11.7%	7.0%	11.0%
N	225	188	66	80	58	61	47	50	30	47